

1/39

FIG. 1A

2736 bp

ATGGCGCCGA GCTCCCGTC ACCCGCCGC CTTACACGG TGTCTGGGG  
 51 GAAGCGCGCC GCCAAGCGCG AGGAGATCCA CCAGAACAAG GAGGAGGAGG  
 101 AGGAGGTGCG GCGCGCGTCT TCGGCCAAGC GCAGCCGCAA GCGCGCATTT  
 151 TCCGGGAGA AGCCCAAGTC GCCCCCCAAG CTGGCCAGCG GGGGAGGAGA  
 201 GAAGAAGGGG GATGCCGAGA TGAAGGAGCC TGCGCGAGGA GGAGCCGAGG  
 251 CCGAGGAGCC CGACGAGGAG GATTTGGGCA TGGCGGGGGG GGTACCCGCG  
 301 GAGCAGGCCA TCCAGGAGGA GTGGTGTGCG TCGCGGGGGG GGTACCCGCG  
 351 GAAGAAGAGG GTGGGAGAAA GGAACGCCGC CGCCGCCGCT GCGGACCACG  
 401 AGCCGGAATT CATCGGCAGC CCTGTTGCCG CGGACGAGGC GCGGTGCAAC  
 451 TGGCCCAAGC GGTACGGCCG CAGCACATGC GCAAGAAGAC CGGATGAGGA  
 501 GGAAGAGCTC AAGGCCAGAT GTCACTACCG GAGCGCTAAG GTGGACAACG  
 551 TCGTCTACTG CCTCGGGGAT GACGTCTATG TCAAGGCTGG AGAAGACGAG  
 601 GCAGATTACA TTGGCCGCAT TACTGAATTT TTTGAGGGGA CTGACCACTG  
 651 TCACTATTTT ACTTGCCGTT GGTCTCTCCG AGCAGAGGAC ACGGTTATCA  
 701 ATTCTTTGGT GTCCATAAGT GTGGATGGCC ACAAGCATGA CCTTGCAGCT  
 751 GTTTTCTTTT CTGAGGAAAA GAACGACAAT GTGCTTGATT GCAATTATCT  
 801 CAGAGTCAAG ATAGTCCATG TTGATCCAAA TTGATGATCA AAAGCCAAGG  
 851 CTCAGCTGAT AGAGAGTTGC GACCTATACT ATGACATGTC TTACTCTGTT  
 901 GCATATTCTA CATTTGCTAA TATCTCGTCT GAAATGGGCG AGTCAGGCAG  
 951 TGATACCGCT TCGGGTATTT CTCTGATGTA TGTGGATCTG GAGAGCTCAT  
 1001 CTAGTATGCC AACGAGGACA GCAACCTTTC TTGATCTGTA TTCTGGCTGT  
 1051 GGGGGCATGT CTACTGGTCT TTGCTTGGGT GCAGCTCTTT CTGGCTTGAA  
 1101 ACTTGAAACT CGATGGGCTG TTGATTTCAA CAGTTTTCGG TGCCAAAGTT  
 1151 TAAAAATATA TCATCCACAG ACTGAGGTGC GAAATGAGAA AGCCGATGAG  
 1201 TTTCTTGCCC TCTTTAAGGA ATGGGCAGTT CTATGCAAAA AATATGTCCA  
 1251 AGATGTGGAT TCAAAATTTG CAAGCTCAGA GGATCAAGCG GATGAGACA  
 1301 GCGCTCTTGA CAGGAGGAAA TTTGTTGTGA AGAAGCTTGT CGGGATATGT  
 1351 TTTGTTGGCA GTACAGGGGA AAATGGCATC TATTTTAAGG TCCAGTGGGA  
 1401 AGGATACGGC CCTGAGGAGG ATACATGGGA ACCGATTGAT AACTTGAATG  
 1451 ACTGCCCGCA GAAATTTAGA GAAATTTGAC AAGAAGGGCA CAAAGAAAG  
 1501 ATTCTCCAC TGCTGGTGA TGTGTATGTC ATTTTGGAG GCCCACCATG  
 1551 CCAAGGTATC AGTGGGTTTA ATCGGTACAG AAACCGTGAT GAGCCACTCA  
 1601 AAGATGAGAA AAACAACAAA ATGGTGACTT TCATGGATAT TGTGGCGTAC  
 1651 TTGAAGCCCA AGTATGTCTC CATGAAAAAT GTGCTGGACA TACTCAAAAT  
 1701 TGCGGATGGT TACCTAGGAA AATATGCTTT GAGGTGCTTT GTTGCTATGA  
 1751 AGTACCAAGC GCGGTGTGGA ATGATGTTGG CTGGTGTGTA TGGTCTGCCA  
 1801 CAGTTCAGGA TGGCTGTGTT CCTCTGGGAT GCTCTTCTCT CCAATGGTCT  
 1851 CCTTAAGTAT CCTCTGCCA CTTATAGTGT TGTAGTACGT GGAGGAGCCC  
 1901 CTAAAGGCTT TTCCGAATGT ATGGTTGCAT ATGACGAGAC ACAAAACCA  
 1951 TCCCTGAAAA AAGCCTTGCT TCTTGGCGAT GCAATTCAG ATTTACCAAA  
 2001 GGTTCAAAAA CACCAAGCTA ACGATGTGAT GGAGTATGGT GGTTCGCCCA  
 2051 AGACCGAATT CCAGCGCTAC ATTCGACTCA GTCTAAAGA CATGTTGGAT  
 2101 TGGTCTTTCG GTGAGGGGGC TGGTCAGAT GAAGGCAAGC TCTTGGATCA  
 2151 CCAGCCTTTA CGCTTTAACA ACGATGATTA TGAGCGGGTT CAACAGATT  
 2201 CTGTCAAGAA GGGAGCCCAAC TTCCCGGACC TAAAGGGCGT GAGGTTTGGA  
 2251 GCAAAACAATA TTGTTAGTGT GGATCCAGAA ATCGAGCGTG TGAACCTTC  
 2301 ATCTGGGAAA CCACTGTTTC CTGACTATGC AATGTCATTC ATCAAGGCA  
 2351 AATCACTCAA GCGGTTTGGG GCGCTGTGGT GGGACGAGG AGTTCTACA  
 2401 GTTGTACACA GAGCAGAGCC GGCATACAGC GTTATAATTC ATCCAGTCA  
 2451 AGCAAGGCTC CTCACTATCC GGGAGAACGC AAGGTTACAG GGTTCGCCG  
 2501 ATTACTACCG ATTGTTTGGC CCGATCAAGG AGAAGTACAT TCAAGTCGGG  
 2551 AACCGAGTGG TTGTTCTGTT TGGCCGCGCA CTGGGCTACT GTCTGGGGCA  
 2601 AGCCTACCTG GGTGAATCTG AGGGGAGTGA CCCTCTGTAC CAGCTGCTTC  
 2651 CAAATTCAC GTCTGTTGGA GAGCGACTGT CGGGGAGGCG GAGGCTCTCT  
 2701 CCGTGTGGCA CCGTTCAGG GAGGATGATT GAGCAG

2/39  
FIG. 1B

1 AGAGCAGCAG CAGCTACGCG AGCCCCCTGCC ATGCGCCGGA GTCGCCCGTC  
 51 ACCCGCCGCG CCTACACGCG TCTCTGGGCG GAAGCGCGCC GCCAAGGCGC  
 101 AGGAGATCCA CCAGAAACAG GAGGAGGAGG AGGAGGTCCG GCGGCGTCC  
 151 TCCGCAACG GCGAGCCGAA GCGCGCATCT TCCGGGAAGA AGCCCAAGTC  
 201 GCGCCCAAG CAGGCCAAGC CGGGGAGGAA GAAAGAGGGG GATGCCGAGA  
 251 TGAAGGAGCC CGTGGAGGAC GACGTGTGCG CCAGSAGGCC CGACGAGGAC  
 301 GAGTTGGCCA TGGGCGAGGA GGAGGCCGAG GAGCAGGCCA TGCAGGAGGA  
 351 GGTGGTTGGG GTCCGCGCGG GGTCAACCGG GAAGAAGAGG GTGGGAGGAA  
 401 GGAACGCGCG CGCGCCGCGT GGCAGCCAGC AGCCGAGATT CATCGGCACG  
 451 CCTGTTGCGG CGGACGAGGC GCGCAGCAAC TGSCCCAAGC GCTACGCGCG  
 501 CAGCACTGCC GCAAAAGAAAC CGGATGAGGA GGAAGAGCTC AAGGCGCAGT  
 551 GTCACTACCG GAGCGCTAAG GTGGACAACG TCCTCTACTG CCTCGGGGAT  
 601 GAGCTCTATG TCAAGGCTGG AGAAAAAGAG GCAGATTACA TTGGCCGCGT  
 651 TACTGAATTT TTTGAGGGGA CTGACCAAGT TCACTATTTT ACTTCCCGAT  
 701 GGTCTTCCG AGCAGAGGAC ACGGTTATCA ATTCTTTGGT GTCCATTAAT  
 751 GTGGATGGCC ACAAGCATGA CCTAGACGT GTTTTTCTTT CTGAGGAAAA  
 801 GAACGACAAAT GTGCTTGATT GCATTATCTC CAAGGTCAGG ATAGTTCAGT  
 851 TTGATCCAAA TATGGATCCA AAAGCCAAAG CTCAGCTGAT AGAGAGTTGC  
 901 GACCTATACT ATGACATGTC TTACTCTGTT GCATATTCTA CATTTGCTAA  
 951 TATCTCGTCT GAAATGGGC AGTCAGGCGAG TGATACCCCT TCGGGTATT  
 1001 CTCTTGATGA TGTGGATCTG GAGACGTCAI CATGATAGCC AACGAGGACA  
 1051 GCAACCCCTC TTGATCTGTA TTCTGGCTGT GGGGGCATGT CTACTGTGCT  
 1101 TTGCTTGGGT CGAGCTCTTT CTGGCTTGAA ACTTGAAACT CGATGGGCTC  
 1151 TTGATTTCAA CAGTTTTCGG TGCCAAAGTT TAAAAATATA TCATCCACAG  
 1201 ACTGAGGTGC GAAATGAGAA AGCCGATGAG TTCTTGCCCC TCCTTAAGGA  
 1251 ATGGCGAGTT CTATGCAAAA AATATGTCCA AGATGTGAT TCAAAATTAG  
 1301 CAAGCTCAGA GGAATCAAGC GATGAAGACA GCGCTCTGA CAAAGACGAA  
 1351 TTTGTTGTAG AGAAGCTTGT CGGGATATGT TATGTTGCA GTGACAGGGA  
 1401 AAATGGCATC TATTTTAAGG TCCAGTGGGA AGGATACGCG CCTGAGGAGG  
 1451 ATACATGGGA ACCGATTGAT AACCTTGAGT ACTGCCCGCA GAAAATTAGA  
 1501 GAATTTGTAC AAGAAAGGCA CAAAAGAAAG ATTTCTCCAC TGCTGTGTGA  
 1551 TGTGTATGTC ATTTGTGGAG GCCCACCATG CCAAGGTATC AGTGGGTTTA  
 1601 ATCGGTACAG AAACCGTGAT GAGCCACTCA AAGATGAGAA AAACAAAACA  
 1651 AGGTGACTTT TCATGGATAT TGTGGCGTAT TTGAAGCCCA AGTATGTTCT  
 1701 CATGGAAAT GTGGTGGACA TACTCAATTT TGCGGATGGT TACCTAGGAA  
 1751 AATATGCTTT GAGCTGCTTT GTTGCTATGA AGTACCAAGC GCGGCTTGA  
 1801 ATATGTTGGG CTGGTTGCTA TGGTCTGCCA CAGTTCAGGA TGCGTGTGTT  
 1851 CCTCTGGGTT GCTCTTCTT CCATGGTGGT CCTTAAGTAT CCTCTGCCA  
 1901 CCTATGATGT TGTAGTACGT GGAGGAGGCC CTAATGCTTT TTCGCAATGT  
 1951 ATGGTTGCAAT ATGACGAGAC ACAAAAACCA TCCCTGAAAA AGCGCTTGCT  
 2001 TCTTGGCGAT GCAATTTACG ATTTACCAAA GGTTCAAAAT CACCAAGCTA  
 2051 ACGATGTGAT GGAATATGGT GTTTCGCCCA AGACCGAATT CCAGCGCTAC  
 2101 ATTCGACTCA GTCTAAAGA CATGTTGGAT TGGTCTTCG GTGAGGGGCG  
 2151 TGGTCCAGAT GAAGGCCAAG CTCCTGGATCA CAGGCTTTA CGGCTTAAAC  
 2201 ACGATGATTA TGAGCGGGTT CAACAGATTG CTGTCAAGAA GGAAGCCAC  
 2251 TTCCGCGACC TAAAGGCGGT GAGGGTTGGA GCAAAAGATA TTGTTGAGTG  
 2301 GGATCCAGAA ATCGAGCGTG TGAATCTGTT ATCTGGGAAA CCACTGTGTT  
 2351 CTGACTATGC AATGTCATTC ATCAGGGCT AATCACTCAA GCGCTTGGG  
 2401 CGCTCTGGT GGGACGAGAC AGTTCTTAOA GTTCTAACCA GAGCAGAGCC  
 2451 TCACAGCCAG GTTATATATC ATCCGACTCA AGCAAGGGTC CTCACTATCC  
 2501 GGAAGAACGC AAGGTTACAG GCGTTCGCG ATTAATACCG ATTGTTGGG  
 2551 CCAATCAGG AAGAGTACAT TCAAGTGGG AACGAGTGG CTGTCCCTGT  
 2601 TGCCCGGACA CTGGGCTACT GTCTGCGGCA AGCTACCTG GGTGAATCTG  
 2651 AGGGAGTGA CCGCTCTCAC CAGCTGCCCT CAGGTTTAC CTCTGTGGA  
 2701 GAGCGCACTG CGGGGCAAGC GAGGGCTCTC CCGTGGGCA CCGCTGCAAG  
 2751 GGAGGTAGTT GAGCAGTAAA AGGATGACAG ATCTGAGCTG AGCTGG

3/39

FIG. 2A

912 amino acids

```

1  MAFSSPSPPAA PTRVSGRKFA AKAEETINQNK EEEEEVAAS SAKRSRKAAS
51  SGKKPKSPPK QAKPGRKKKG DAEMKEPVED DVCAEEPDEE ELAMGEEAE
101 EQAMQEEVVA VAAGSPGKKR VGRRNAAAAA GDHEPEFIGS PVAADEARSN
151 WPKRYGRSTA AKKPDEEEEL KARCHYRSK VDNVYVCLGD DVYVKAGENE
201 ADYIGRITEF FEGTDQCHYF TCRWFFRAED TVINSLSVIS VDGHKHDP RR
251 VFLSEEKNDN VLDCTIISKVK IVHVDPNMDP KAKAOLIESC DLYYDMSYSV
301 AYSTFANISS ENGQSGSDTA SGISSDDVDL ETSSSMPTRT ATLLDLYSGC
351 GGMSTGLCLG AALSGLKLET RWAYDFNSFA CQSLKYNHPQ TEVRNEKADE
401 FLALLKEWAV LCKKYVQVDV SNLASSEDQA DEDSLDKDE FVVEKLVGIC
451 YGGSRENGI YFKVQWEGYG PEEETWEPID NLSDCPQKIR EFVQEGHKRK
501 ILPLPGDVVD ICGGPPCQGI SGFNRYNRND EPLKDEKNKQ MVTFMDIVAY
551 LKPKYVLMEN VVDILKFADG YLGKYALSCL VAMKYQARLG MMVAGCYGLP
601 QFRMRVFLWG ALSMVLPHY PLPTYDVVVR GGAPNAFSQC MVAYDETKPF
651 SLKKALLLGD AISDLPKVQN HQPNVMEYGG GSPKTEFQRY IRLSRKDMLD
701 WSFGEGAGPD EGKLLDHQPL RLNNDDYERV QQIPVKKGAN FRDLKGVVRVG
751 ANNIVEWDPE IERVKLSSGK PLVPDYAMSF IKGKSLKPFQ RLWDETVPT
801 VVTRAEPHNQ VLIHPTQARV LTIRENARLQ GFPPDYRLFG PIKEKYIQVG
851 NAVAVPVARA LGYCLGQAYL GESEGSDFLY QLPPSFYSVG GRTAQARAS
901 PVGTPAGEVV EQ

```

4/39

FIG. 2B

RAAAAATAAPAGAPSSPSPAAPTRVSGRKRAAKAEIHHQNKEEEEVAAAS  
 SAKRSRKAASSGKKPKSPFKQAKFGRKKKGDAEMKEPVEDDVCAEEPDEE  
 ELAMGEEAEQEQAEVVAAGSPGKKRVGRNAAAAAGDHEPFTGS  
 PVAADARSNWPKRYGRSTAAKFPDEEELKARCHYRSKVDNVVYCLGD  
 DVYVAGENEADYIGRITFEFGTDQCHYFTCRWFFRAEDTVINSIVSIS  
 VDGHKIDFRFVFLSEKNDNVLDCCIISKVKIVHVPNMOPKAKAQLIESC  
 DLYYDMSYSVAYSTFANISSENGQSGSDTASGISDDVDLETSSSMPTRT  
 ATLLDLYSGCGGXTGLCLGAALSGLKLETRKAVDFNSFACQSLKYNHPQ  
 TEVRNEKADEFLLKEWAVLCKKYVQVDVSNLASEDQADESPDKDE  
 FVVEKLVGICYGGSRENGIYFKVQWEGYGPEDTWEPIDNLSDCFQKIR  
 EFVQEGHKRKLPLPGDVDVICGGPPCQGISGENRYNRDEPLKDEKNKQ  
 MVTFMIDIVAYLKPXYVLMENVVDILKFADGYLGKYLALSLVAMKYQARLG  
 MMVAGCYGLPQFRMRVFLWGAALSSMLPKYPLPTYDVVVRGAPNAFSQC  
 MVAYDETQKPLKICALLGDAISDLPKVQNHQPNVMEYGGSPKTEFQRY  
 IRLSRKDMLDWFSFEGAGPDEGKLLDHQPLRLNNDYERVQQIPVKKGAN  
 FRDLKGVVRGANNIVEWDPETIERVKLSGSKPLVPDYAMSPFKGKSLKPPG  
 RLWDETVPTVVAEPHNQVIHPTQARVLTIRENARLQGFDDYTRLFG  
 PIKENYIQVGNVAVPVARALGYCLGQAYLGESESDPLYLQLPFSFTSVG  
 GRTAGQARASPVGTPAGEVVEQ\*KKDRSELSEW

5/39

FIG. 3

Primer	Sequence 5' - 3'
1F	TGGTTGCTATGGTCTGCCACAGTTCAG
1R	CCAGCTCAGCTCAGATCTGTATCCTTT
Seq2FN	CGAAAGCTAATCTACACAACAGC
Seq2RN	GATCCTCTGAGCTTGCTAAATTTG
3R	CTCATCTTGGAGTGGCTCATCAC
S3F	GAGCACATGAGGGAGAGTGTG
S3R	TCTCTAATTTTCTGCGGGCAG
4F	CCTCTGCCACCTATGATGTTGA
5F	TAAAGGGCGTGAGGGTTGGA
7F	TCACATTTGTATGGCAGGTTATC
8eF	CTGAGGAAAAGAACGACAATGTGC
8eR	GCAATCAAGCACATTTGTCGTTCTTTTCCTC
9eF	GAAGAAGAGGGTGGGAGAGAAGGACG
9eR	TTCTTTGCGGCAGTGCTGCG
11fF	GTATTGAATTGATTTCTCAACTAGTGCAC
11fR	CAGGCTCAACGGCGATG
12fF	TATGCTTTCATCACATAGACCCAAAGTC
12fR	GATAGACCTAATGCCAAATGAGATTAAG
13fF	GCGATCTTCAGTCTCCACCATC
13fR	GAAGACGTGCCTCCATGTTTCATC
14F	GTTGGTTCTTCCGAGCAGAGG
14R	GACTGCCACATATCTTATTAATCGC
15F	GCATGTGTGACGAAATGCTTACATTC
15R	CCTCTGCTCGGAAGAACCAAC
16F	CTGTTCCGAGATTTCATGCATGATG
16R	GGAGAACAGAATGGTTGATTCAATGG
17F	GCACTTCACTCTCTGGCAAACC
17R	CGGTACGCTGCTGCTGCTCTC
18F	CCATAGCATCTCACATATCGCAAGG
18R	GGAAAGAAGGCAGTTAGTTGTAAATGGG
MU	AGAGAAGCCAACGCCAWCGCCTCYATTTTCGTC
RaceRT	CTACAACATCATAGTTGGGCAGAGG
AP2 marathon	ACTCACTATAGGGCTCGAGCGGC
T7	TAATACGACTCACTATAGGG
Sp6	GATTTAGGTGACACTATAG
M13F	GTTTTCCAGTCACTGAC
M13R	CAGGAAACAGCTATGAC

6/39

FIG. 4

Gene Name	Organism	Function	General Structure
<i>DNMT1/hnmt1</i>	human/mouse	maintenance	
<i>ME1</i>	<i>Arabidopsis</i>	maintenance	
<i>Zmct1</i>	maize	putative maintenance	
<i>DNMT3/hnmt3</i>	human/mouse	<i>de novo</i>	
<i>Zmct3</i>	maize	putative <i>de novo</i>	
<i>DRM</i>	<i>Arabidopsis</i>	putative <i>de novo</i>	
<i>CM11</i>	<i>Arabidopsis</i>	undetermined (putative CpNpG)	
<i>Zmct2a</i>	maize	CpNpG (maintenance and/or <i>de novo</i> )	

Figure 5

**SUBSTITUTE SHEET (RULE 26)**

cgtgaggagcagcgtgtgcgcgaggagccgcacyaggaggagtgtggccatgggcgaggaggaggccgaggagca base pairs  
gcacgtcctcgtcgcacacgcggctcctcgggctgctcctcctcaacgggtaccgcctcctcctcgggctcctcgt 1051 to 1125

HaeIII HpaII 96F  
ggccatcgaggaggaggtgttcgggtcgcgctggggtcacccgggaagagagcggcgggaagaaacggcgc base pairs  
ccggtacgtctctctctcaccaaagccacgcccgcgcccatggggccctctctctccaccctctctcttcggcgcg 1126 to 1200  
ScrFI  
SmaI

cgccgccgctggcgaccacgagccggaagtctcatcggcagccctgttgccgcggacgaggcgcgacgaactggcc base pairs  
gcgggcggcgaccgctgggtgctgcggcctcaagtgcgcctcgggacaacggcgccctgctccgcgcgtcgtgaccgg 1201 to 1275

c aaagcgctacggcgcgagcacttgcgcgaagaagtcacattatctttcccgagctctgggttttgattcgacca base pairs  
g tttcgcgatgccccgctgcgcgaaggccttttcccatgtataaaagagggctcgagaccaaaactaaactggt 1276 to 1350

9eR

gattctttactccatgtctgtctagtacttgcgagctgagcaatctgctattctgtgatttattgtgcgtgcagacc base pairs  
ctaaaaatgaggtacagacaatcatgaaecgtctgactcgttagacgataaacgactaaataaacgcgcgtctcg 1351 to 1425

ggatgaggaggaaagagctcaagccagatgtctactaccggagcgctaaaggtggacaaagtcgtctactgctctcg base pairs  
cctactcctcctctcagatccgggtctacagtgatggcctcgcgattccacctgttcgacgagatgacggagcc 1426 to 1500

MspI

ggatgacgtctatgtcaaggctcttctgttcacgcgtttctgttgcttctgctctcatttatgatgcgcataatgtgt base pairs  
cctactgcagatacagttccaggaaacagtagcgaaagacaacgaagacgagagtaaatatactacacgtatacaca 1501 to 1575

AvaII

ttgttaaggaagcaagaattgcttgattttgtgtgcgactcgcatttcctgtgacgagttctcggtatggteacc base pairs  
aacaattccttcgttcttaacgaactaaaaacaacggctgagcgtaaaggcactgctcaagacgcataccagctgg 1576 to 1650

TaqI BstNI      Sau3AI  
 ggtacgtggcactgatacacacgtggtatgctggaagtctggtagcatatatttgcacgcaccaggagggtccaga base pairs  
 ccacgcaccgtgactatgctgcaccatacgcaccttcagaccatcatataaaacgtagctgggtctccaggct 1651 to 1725  
 EcoRII AvalI

tgcgatatgtgcggtatagtgcttatttgattgcacccgttccgagatccatgcacatggcggtgcttagatgac base pairs  
agctatcacacgccatatcacgaataaaactaacgtgggacaagcctcttaagtacgtaactccgcacaaatctactg 1726 to 1800

PvuII EcoRII PvuII HaeIII HpaII NheI HinfI  
gcctccagacagctgcctgcaggcagctgattctgcggcaggcgtcggaatgggtgaagtgtgcgtgcgaaga base pairs  
cgaagctctctcgcagcaagctcctgcgataaagacgggctcgcagccttaccactcgaacgcgcagctct 1801 to 1875

ScrFI

aagagtcCGgtggatgggtttatcgggaGctCgtataacgtacgaagaaaaaaacaagagaaaggaagatataaa 1876 to 1950  
BstNI

tacaqtaacaaatcacttcaaagtgtaacgtgcacagtaccttataaaatgaaagtttagttgctccctctacgacg 1951 to 2025



9/39

FIG. 5

Continued

EcoRV      Continued

attgagggtgtgtgataattattacatactagaagatatcgatggtgccattgggaattgcgaagaatgtggaa base pairs  
taactccacacactattaataatgatgatcttctatagcagctacaacggtaaccttaacgctcttaccctt 2026 to 2100

MseI      MseI

atgatgtrggtagccttgattaaagaggttaacagtgtagtgggatgacatgactattagttagagatgatgtggat base pairs  
taccacaaccatcgaaacataattctccaattgtgaatcacccctactgtactgataatcaatcctctactacacct 2101 to 2175

agtaagtgggatatgatgtagatgacttgtgtgttgagacagaactataacatggagtggaaatgggagcagca base pairs  
tcatttcacctatactacatctactgaacacacaactctgtcttgatatattgtacctcaacctttacctctgtcgt 2176 to 2250

MseI

tgggtcaaacataccctaaatgctgtctctacacaatgtgggtgattgggtgtatagtctgggtgttaaaagctggat base pairs  
accaqgtttgtatgggatttacggacagagatgtgttacaccactaaccacatatcagaccacaattttcgacct 2251 to 2325

HinfI

**MseI**

Xba I

Scrf1

EcoRII

actttgattctgttgaagattgtcacaccggaaatttaaggacaaatctagatacatctcatatgtgcaccggat base pairs  
tgaaacraaagacaacttctcaacagtgtggccttaaatcctgtttacatctatgtgaagatatacacgtggtccta 2326 to 2400

**BSENI**

agtgtatagataccaatgtcataatctttattacacgcagataatgtcttcaaaaatctgggtgttacaagatg base pairs  
tcacatatctatggttacagtattagaaaatagtgcctgctattacagaatgttttatagaccacaatgttctac 2401 to 2475

Msei

Msei

cacccttccaacatggttaatgctgcgaactgttttaattaaacagaatgcagtgctttgaacaaaaaaatgctgc base pairs  
 gtggaagattgtacaaattacgacgtttgacaaaaattaatttctctacgtcacaaaactgtttttttacgacg 2476 to 2550

MaeI

154 F

Sau 3A:

HinfI

MseI

tttatcctgcactcttctgttttcgcatgctcagcaaatctgcttaccatccattatgatctctgagatctctttaaattt base pairs  
aaataaqaacgtagaaacaaacgtacacacagctcttaacgaatgttaacggttaatactagagactctaagaattttaa 2551 to 2625

ctagcatgatgaaagtatttactaatccaactgaacacaaaacattgtttgaatgaacaaggcaacacggatgctt base pairs  
gacgcgtactactttcataaatgattaagttgactttgtgtttgtaacaaacttactgttccgttgtgcctacgaa 2626 to 2700

MseI

ggaataaagggtgtgtataaatcacttagtgggttttgcctctcacaccacatcttcatgggttctttaaataa base pairs  
ccctattaccaaacacatattatagtgaatcaccaaacgaagagtgtggtgtagaaagtaccaagaaattattat 2701 to 2775

MscI

HaeIII

ggtactgactcttaagttcttattcctttttgtctatcttagctggagaaaaacgaggcagattacattggccgc base pairs  
caatgactcgaattcaaaagataaggaaaaacagatagaatcgacctcttttgcctcgtctaattgaaccggcg 2776 to 2850

14eF

actactgaatttttggagggaactgaccagtgctactattttacttgccgttggttcttccgagcagaggacacg base pairs  
taatgactctaaaaaactcccctgactggtcacagtgtataaaatgaacgggaaccaagaaggctcgtctcctgtgc 2851 to 2925

15eR

ggtgtgatttagtattttgtcatctcatgcatggtggattttctggaaatgtggaaaaacacagcactctctc base pairs  
cacacataaatctataaaacagtaaaqatcgtacacacctaaaaagaccttacaccttttqtatctcgtgagagag 2926 to 3000

MseI

HaeIII

HaeIII

tacaccacacacacttctagtatatgtgtacacgtaattggggccaacactagacacatggcccaacatccccct base pairs  
atgtgtgtgtgtggaagatcatatcacatgtgcgaattaccgggtgtgtacttgttaccgggtgtataggggga 3001 to 3075

EcoRV

caagatggcgatagatatcaatcatcccatcttgctacataacacatcacactcttttactcctataacctta base pairs

10/39  
FIG. 5  
Continued

gtctacaccgctatctatagtttagtaggggtagaacgagttattgtgtagtcgagaaaaaggagataggggaat 3076 to 3150

HinfI

gccaagcaatctgctatttgaccccttttgagctttacatgattcaactctaaagaccattatctaactctctcttg base pairs  
cagttcgttagacgataaaactggaanaactcaaatgtactaagttgagatttcattggtcaatagattgaagagaaaac 3151 to 3225

ClaiI

atgaagaatcgatcaattttccacatgtttttgtctatcatgttgaacggattgttagctatattcatggtgcac base pairs  
tacttcttagctagtttaaggtgtacaaaacagatagtcacaacttgacctcaacatcgatataagattaccgactg 3226 to 3300

TaqI

Sau3AI

HinfI

MseI

18eF

ctattatcacaccataacttcaggagagctttttcttaatacattcaactctcgatagaagagaccctttatcccatagc base pairs  
ataaatagctgcggtattgaagtccttcagaaaaaatatgttaagttgagactattctctgggaaatagggtatcg 3301 to 3375

HaeIII

attccacatattcgaaggccatagctcggattctctgtcttcggcggtgaacgggataccacagattgtttcttg base pairs  
tagagtgtagctagcttccgggtatcgagccataaagacgaagcgcgcaccttgccctatgggtcttaacaaagaac 3376 to 3450

cttctccatgatactaaatttctccaaacaacacacaatattcctgaagttgaccttctatcatcaaggcaacta base pairs  
gaagaggtagctatgatttaagagggtgtgtttgtgtgttataggacttcaactggaaagatagttctcgttgat 3451 to 3525

MseI

SerFI

EcoRII

ccccagctctgcatcagagtaaccttccaccttttagatgacctgacctttaaagattattccctttccaggacaa base pairs  
ggggtcagacgtagttctcatctgaaggtggaactctactggtcatggaaatttctaataagggaagggtcctgtt 3526 to 3600

BstNI

TaqI

gtcttcaagattcgcagatacgcatacactgcatcaagattgccacttctggggtcatgcataatcgactacc base pairs  
cagaagttcatagcgtcatatgctatgtgacgttagttctacaggtgaagacccagtaacytatatagctgagtgg 3601 to 3675

HinfI

EcoRV

acactgactgcataatgtatagctcaggtctgttatggcacaagtagatgagccgtccaaagcttttgcatacctt base pairs  
tgtgactgacgtatacacattatgctcagaaacataccgtgttctactctggcaggtgttcagaaactatgaa 3676 to 3750

Sau3AI

HinfI

HinfI

TaqI

HaeIII

tctttatccacaggtcacccagattcagcacataatttatgattcaagtcgtaggtgtgtctacagggccgacac base pairs  
agaaataagtgctcctagtggtctaaagtcgttataaataactaaagtcagctatcccaacagatgtcccggtctg 3751 to 3825

Sau3AI

Sau3AI

cccaacatacctgtttcatcaagtagattcaaaactatttctcttgggagagaactattccttttggagatcga base pairs  
gggtgtatggacaagtagttctcatagattttgtataaaggaaccctctcttgataagggaacactctagct 3826 to 3900

BglII

TaqI

Sau3AI

HinfI

gcactctcaataccaagaagattttgagatgaccaagattcttaaacctcaaatcttacttagattctctttt base pairs  
cgttagagttatggttctttcataaactctactggtttctagaattggagtttaagggaatgaactcaagaagaaa 3901 to 3975

BglII MseI

Sau3AI

agacatgcgaattcgaagtcggcatcactgttaataataatcatccacatcacagctagaattgcgaattcgt base pairs  
tctgcaagttagagttctagccgttagtggaactattattatagtaggtgtatgtgtcgatcttaacgttaagca 3976 to 4050

Sau3AI

cgcccaagttgtgataaaaaacagctgtgactcctcgttcgattgtttatattcccatgctcattgtgcagctcta base pairs  
gcaggtttccacaactattttttgtcacactagaggcaacgttaacaaatatagggtacgattgataacgtgcagat 4051 to 4125

TaqI

aattctgtcaaacctgctctcttggggactgcttgagaccatacaatgatttttcaatcgacaaactttcccaatt base pairs

11/39  
FIG. 5  
Continued

ttagacaggtctggtagcagaacccctgacgaacctcggtagcttactaaaaaagtttagctgttgaagggttaa 4126 to 4200

ScrFI

EcoRII Sau3AI

gtctcaggtcttgacaatccaggaggatctccatatagaccctctctgcacaaatcaccattatgaagaagcattc base pairs  
cagagctccgaacctggttaggtctctccatagaggtatattcggaggagaacgttttagcgttatcatctttctgtag 4201 to 4275  
BstNI

MseI

HaeIII

Sau3AI

ttaacatcttagttgatacaaggccatccaaaatttgcagcacagaagatcaatgctcttacagtactcatcttt base pairs  
aatgttagatcaactatgttccccgtagggttttaaacgtcgtgtctctcagttacaggaatgctcatgagtaaaaa 4276 to 4350

gccactggtgcacaaatgctctatcataatcaattccatatgtttgactataccctcttgcaaacatcttctgtctta base pairs  
cggtagaccacgtttcacagagtagtatagrtaaaggtatatacaaaactgatatgggagaacgttggtagagaacaaat 4351 to 4425

tatcgtttctacccttctctctgggttttgccttccacagtgataccctttacaactaacgtccttcttctcttta base pairs  
atagcagaatgggaaggaagacccaaaacgaaggtcacttatgggtaaatgttcaatgacggaaaggaagaaat 4426 to 4500

161R

XbaI

MseI

ggtagttttccaaattcccaagtttgattttttctagagcttttaagctctctccaaactatgcttcacgcagctta base pairs  
ccatcaaaagagttaaagggttcaaaactaaaaaagatctcgaattcggaggaggttgtaaacggagtcgggtcaat 4501 to 4575

gaattacattgtgctctcttccaaatctcttggaaattgctacggaatgcaatgtagcaaaaatgctctatatgat base pairs  
cttaagtgaacacgaagaagggttagagaaccttaacgatgcttactgtactactcgttggttacagagatatcta 4576 to 4650

HinfI

ggtagacaaagacgcatatgagacataattgctaatgctcatgttcatatccatcccttgttggggggagctccagct base pairs  
ccactgtcttctgctatactctgtattaaacgtatcacgtacaaagtattaggtatggaaacacccccctaggctcga 4651 to 4725

HhaI

ttagcagcgcctccttctcgtattgcaatgggcaaatcataagtgtcataatcttccagtcttctccatgagagctc base pairs  
aatcgtgcgcagggaagaagcataacgttaccctgttagtattcacagattagaagtcaaaagggtcactctgcag 4726 to 4800

aaaggtagatttatagcctcttaattgctgtttggagagaactgctcagtagctgctgaattgggtttccagagcc base pairs  
tttccacgtcaaatatcggagattacacaaacctctcttgacgagtcargaaactcagacttaaccaaagctctctcg 4801 to 4875

tgagggtgcacatgggaactttctcttgatataacttcgcccctatatcgttaagtgctcccaagagatttatta base pairs  
actccaacgtgtacccctgaagaagaacatatagaagcgggaatatagcattcagcagaggtgtcttcaataat 4876 to 4950

ttctcgtgactaggatgtgtctccaattcacttggcattacttgcattcttttgagaagaccacatcaccacttcc base pairs  
aagagcactgtctctacacagaggttaagtgaacccgtaatgaacgtagaaaaactcttctcggttagtggtgaagg 4951 to 5025

HinfI

TaqI

attttatttgggtgttgttccattgaaatcaaccattctgttcttccccctctcgactagcttctatctgtgctagta base pairs  
tcaaatcaacacaaacaaaggaatctttagtttagtgaagacagaaggggagagctgatcgaggtagacacgactcat 5026 to 5100

161R

HinfI

Sau3AI

gagacagaatcaagaaaaaaatttagatctgtcttctcaccatagaagagccacagctctctctaaatgtaacatcc base pairs  
ctctgtcttagttctcttttttaaatctctagacagaagagtggtatcttctcgtctcagagagattacattgtagg 5101 to 5175

BglII

HinfI

PstI

atgcttcaaaaacaaacgtcgttctcactaggactcccaacttgtatcccttttgccttcgaggtatctcaacaaaa base pairs  
tacgaattgttctgttcagcagatgactcctgaggttctgaacatagggaacacgggacgctctctataggttgcttt 5176 to 5250  
EcoRV

BamHI

Sau3AI

atgcacttcacagcagaggatccaacttccccaccctgaggcttatgtatctctgcacaaaactgtacattcaaaa base pairs  
tacgtgaaggtcgtgctccttaggttgaagggtggactccagatactagagactgttttgacatgtagggtttt 5251 to 5325

Sau3AI

12/39

FIG. 5

Continued

HinfI

HinfI

atTTttaggtggaaaccacaaacttattctcaccgagaagaatctcacatggagtccttcattgcagtagttttttaa base pairs  
 taaaatcccaacttggtgttgaataagatggctctctcttagagtgtagctcagaagtaacgttcataaaaactt 5326 to 5400

MseI

ggagtcgattataaagatatgtggcagtcacatcagcttcactccataggaacttcgggaacattctgtgtaaac base pairs  
 cctcagcgaatatttctatcacacccgagcttatgtcgaagtggagtagtccctggaagcctgtgaagtaacatttg 5401 to 5475

141R

HinfI

atcagcgaacgagcaacttccaaaatgtgcagattctctctttcagccacaccattttgtggagtgtagtcagg base pairs  
 tagtcgttctgtcgttgaaggttttcaactgttaagaaggaaagtcgggtgtggttaaacacctccacatagtcct 5476 to 5550

MseI

caggtagtctgatgtaatataccattctcttcagacagaatgcattaaatccctgtttacatagctgggtccattg base pairs  
 gtcttcagactacattatattggtaaagaactgtcttttaagtaatttagggaaacaaatgtatgagccaaggtaac 5551 to 5625

111F

HinfI

tctggctcttaggattttgacttgagttatggaatgaatctcgaactagtcgcacaaaaattttgaaaacatttcaat base pairs  
 agaccagaatcctaaaactggaactcataacttaactaagagttgatcacgtgtgtttttaaacttttgtggaagta 5626 to 5700

121F

TaqI

acttcatctttatgcttcatcacatagacccaagtcattccgagaaaaaacatcgataaagtaacaagcatttc base pairs  
 tgaagtagaaaatacgaagtagtgatctgggttcagtaaggctcttttctgttagctatttcattgtttcatgaag 5701 to 5775

ClaI

MseI

atcccatataatagaagtcacaggacatgtccaaacatcagaatgaactagccaaaaggagataatactcctgata base pairs  
 tagggtaattatcttcagtgctcctgtacaggttctgtagcttcaactgatcgtgtttctctctatatagaggactat 5776 to 5850

TaqI

cctcgactaataaagatgtccttctgtgtgttttgcacaaactcacaggcatcacacaatagctgtctttatccacc base pairs  
 ggagctgattatattctacaggaacacacaaaacgtttgagtgctcagtagtggtttatcgaacgaaaatagggtgg 5851 to 5925

HindIII

ccactcattacatcaggaagaaagcttgcatacttatacaaaagaagatgcacctaatctacaatgcaagagcatc base pairs  
 ggtgagtaatgtagctccttttcgaaacgtatagaatagttttctcttcacgggattagatggtacgttctcgtag 5926 to 6000

Sau3AI

actgcaacctctctctcttccattcttctgtgccagcatagtgcatattgtaccatttagtccctcagcatccata base pairs  
 tgagcttgggaggaagagaaggtgaagaacaacgggtcgtatcagtgataacatggtaacaggggagtagctaggtat 6001 to 6075

SrfPI

MseI

taccacaatccattacgctcgttagctgtcccaagctcttccctgtttccctctcctgaattaaacaattatct base pairs  
 atgggttagtagtaatgcggaccctcgacaggggttcagagaagggacaaagggagaggacttaattgttaataga 6076 to 6150

BstNI

TaqI

Sau3AI

EcoRV

cgatcagaataatacgaacatccaattgatcaaccaaggcacttagtgatatcaagtgcaggaanaaggtggc base pairs  
 gctagtcttattatctgtcttaactagttgggtccgtgaactcactatagttcaactgtccttcccaacgc 6151 to 6225

Sau3AI

MseI

acatacaaaactgatgacaacttaatatagatggagtgcatctgactgtgccaacaccccttgatgggtgttggtgta base pairs  
 tgtatgttttgactactgttgaattatctacctcagtaacgtgacacgggttgggaactaccacaaccacat 6226 to 6300

EcoRV

ccatcagcagtttgtataattcttctacgtgtgggggagattctatatatgatgtaaatctcagtgagctgcct base pairs  
 ggtagtcgtcaaacatataaagaatgcacacccccctatagaatataactacattaaagtgaacctgcacgga 6301 to 6375

13/39

FIG. 5

Continued

HinfI MseI  
 gtgacatgcttttgatgctcctgagctcaaaatccattttaactggtgacccgctggtggtacaaaagcatgagca base pairs  
 cactgtacgaactacgaggactcagatttttagttaaattgacacacgtgacacacccattgttttcgtactcgt 6376 to 6450

HinfI Sau3AI  
 taattaccttcattcagctaggcggaagtgacaaaatccctctgtgtgagactcctgactctttatctccagagatt base pairs  
 attaargggaagtgcacatccgcttcacctgttttagggggacacactctgaggactagaaaatagaggtctctaa 6451 to 6525

MseI  
 tggattttctctcccaactttgtttcatcttcgtgttccataaaatgttccaagtctctctgtgtgattgtgcga base pairs  
 actaaaaagaaggagttgaacaaaagtagaagcacaaggtatttacaagttcaagaagaacacattcaacagcgt 6526 to 6600

MseI  
 ttgcaccttgcccaactcctgctccacgaacctctgccgaccttagagaccctctctctctccacagattaaat base pairs  
 aagcgggaacgggttgaggacggaggtgctggagacggcgagatctctggggagagagagaggtgtctaattga 6601 to 6675  
 ttggaaagcttagaacaattacgtgcaatattgtccaacattaccacaattgtacaactctctctagtattcttggtt base pairs  
 aaccttcggaatcttgttaatgcagcttatcacaggttgtaattggttgaacttgaagagatcatagaaccaa 6676 to 6750

SerFI  
 HinfI HinfI EcoRII  
 ctcatagctgaaaaacagagtagggcggtttgagaactttctctcatcactttgagctcttgactctctctgg base pairs  
 gagtacgcagctttttgtgtcctactcgcgcgaactcttgaaagagagtagtgaacactcagaactcaggagagacc 6751 to 6825  
 BstNI

TaqI  
 gatatggcagctatggcttcttgragctaggaagagtgagtagtgaacaatcgaggacgctcttccctcgcaac base pairs  
 ctataccgtcgatcacgaagaacattccgactctctccactaaatatttctgactctcggtgagaaggagcttg 6826 to 6900  
 131R

MseI  
 tctgagtttagccccccttagcaattgaagtacacgctctttttccacccattctctcgccaagcaacacactct base pairs  
 agactcaaatcgggggatcgttaactctatgtgcagaaaaaagggtgggtaaagaagcgggttcgttgtgtgaga 6901 to 6975

Sau3AI Sau3AI  
 gagggtgtgtagctcaataggatcataatgatcaacatcagcccaataacattgtaactcctgaacgtactccgcc base pairs  
 ctccacaccatcgagttatcttagtattactagttgttagtcgggtatttgaacattgaggacttgcattgaggcgg 6976 to 7050

Sau3AI Sau3AI 131F  
 acagatcgctccccctgtttgatattatggaggcgactcttcagctctccacatcaacataacatttccagctccc base pairs  
 tgtctacgagagggggacaaactataactctccgctagaagctcagaggtcggtgtgtattgtaaaggctcgaggg 7051 to 7125

PstI HinfI  
 gaggtaattttctcaagctcctccacattttctgcagcactttatgattgtatcaacagctgtagcaattgtctga base pairs  
 ctctgtgaaagaggttcacgaagggtgtaagaagctcgtgaatactaacattagttgcagcatcgttaacgacct 7126 to 7200

MseI  
 atcatagaactcaacatccacgctgccactaaaagagtttatagcatcccagctcttccattctcatcattacta base pairs  
 tagtatctttagtctaggtgcgagcgtgattttctcaaatatcgtagggtcagaaggtaagtagtgaattgaat 7201 to 7275

HinfI XbaI  
 MseI XhoI  
 tccctgggctcaacgacatccctttaacatagccctcgagctctctttgccttcaataatcgcaatgctcttcta base pairs  
 aggaacccgagtgctgtgagaggaattgtatcggagctcagagaaacgggaagttattagcgttaccagaaagat 7276 to 7350  
 TaqI 541R

MseI  
 gaccatgcgcaataaattttccacccctcttaacttcaacttggcattaggctctatctctccgaactgctct base pairs  
 ctggatcggtttattaaaaagcggggaagattgaattagagattaaacgtaatcagatagaagagcttgaccaaga 7351 to 7425  
 121R

MseI TaqI  
 atatgagcaacattgtctttaattgatggagcctcatcccttttctgtacagtaattcgaccaatttaca base pairs  
 tatactcgtttgaacgaaattactactacctcggagtaggggaaaaacagctgtcattaaagctgggttaaatggt 7426 to 7500

Pyru I I

PvuII  
gctgc base pairs  
cgacg 7501 to 7575

ScxFl

Sau3AI  
gg base pairs  
cc 7576 to 7650

CCGGC

ac base pairs  
cg 7651 to 7725

Sau3AI

ac base pairs  
cg 7726 to 7800

gc base pairs  
cg 7801 to 7875

ScIFI

cg base pairs  
gc 7876 to 7950

## Haft I

c base pairs  
ag 7951 to 8025

EcoRII

g base pairs  
c 8026 to 8100

HaeIII

ac base pairs  
cg 8101 to 8175

gg base pairs  
cc 8176 to 8250

8eF

aa base pairs  
at 8251 to 8325

Pst

ac base pairs  
cg 8326 to 8400

15/39

FIG. 5

Continued

BamHI

PvuII

ataatgtttctgactaccatgttttggctcctactgctctagatggatccaaagcgaaggctcagctgatg base pairs  
tattacaagactgtagtgtaacaaaacacggatgaacggatctacctagggtttccggctccgagctgactatc 8401 to 8475  
SauIAI

agaggtgcgacctatactatgacatgcttactcctggtgcataattctacatttgctaaatctcgtctggtaatt base pairs  
ttccaacgtggatgatgactgtacagaaagagacaacgtataagatgtaaacgattatagcagacattaa 8476 to 8550

MseI

cttctgcatactcttttttggctgactgctgaatgcagctagctttgccaaggctcaaatcacatgagttgtr base pairs  
ggagagcgtagtagaaaaaaccaactgactgactttacgtcaactcgaaacgggtttctcaatttatgtactcaaa 8551 to 8625

TaqI

MseI

MseI

cctgcactcgaaaaggatgtcaataatgtccacaacctcgaaaatgtatttttagatacttaactgttgtaag base pairs  
ggacgtgagcttttccctacagttattacagggtgttgagacttttaccataaaactctatgaattgaacaattca 8626 to 8700

cagtaaaacctgtcagatacttgggttttgggtacgataccatccttatgtagctaaaacctcgtaagggtatg base pairs  
gtcattttggacagctctatgaaccccaaacctcgtcaatgttaggaatacactcattttgagcagttccctaca 8701 to 8775

TaqI

Seq2FN

caatgacgtgttgattgtgtatcagatattctgtgttcgaaagctcaactcgaacaaacagcttatgtaatga base pairs  
gttactgcacaactaaacacataatccataagcaaaacagctttcgatttagagctgtgttgcgaatacattactat 8776 to 8850

HindIII

HaeIII

aaacctcaaacaaacttgccctttcctaagcttaggtttataggattagcgttttagtgcatgtaaggcctatttg base pairs  
tttggagtttgtttgaacggagaagtattcgaaatccaaatctccttaatcgaaatcacgtacattccggataaac 8851 to 8925

HaeIII

BstNI

ScrFI

TaqI

ScrFI

cttcacggctccctcgccgagctcctggcttagacagccatcctggcgtaggtgccrcaaatcgaaacacctggga base pairs  
gaagtgcggaggaggcggctcgaggacagctctgtcggttaggaacggcatccacgggctttagctttgggacctt 8926 to 9000

EcoRII

BstNI

BstNI

ScrFI

HaeIII

ScrFI

EcoRII

ggccagctttgactagcaggttttctgggtgcaaaccaaacacgcttatagttgcaagtataactgaattggt base pairs  
cggctgcaaacgtgatcgtccaaaggagactcagctttgggttgcgggatacacagcttatattgacttaacca 9001 to 9075

BstNI

MseI

gctcacctttgtctaatagtcttaagtttttgggttttcatcggtgcatgcaactcatactcaatagtcataatga base pairs  
cgagtggaaacagattatcgaaattcaaaaacaaagtagccacgtacgttgaggtatgagttatcagttatcact 9076 to 9150

BstNI

XhoI

HinfI

tatatgtgttcagcatagaactcttcagtttgatctctggcaggggcaatcaataaaataattgcagcttacc base pairs  
atatcacaagttcgtatcttggagagctcaaaacttaggacgtcccggttagttattttataacgtcgaaatggg 9151 to 9225

TaqI

EcoRII

ScrFI

S31P

ctattttcagctttgagcactgagggagagtgtagaattataagtggtttctccatctttctcaacagatgaa base pairs  
gataaagatgcacaactcgtgtacttcccttcacaacttaatttcacacaagaggtagaagaagatgtgtcactt 9226 to 9300

HinfI

MseI

MseI

ctgggtttgctgtaactcaatgatatttgagtcaaatgttttacttttaaatcatagtgatgcaatttaact base pairs  
gcccaaacacgtacatttggtttactataaaactcagtttacaagaatatttagtatacactacgttgaatta 9301 to 9375

16/39  
FIG. 5  
Continued

aacatattttttggtctcgtgtgagggagtgtagctataactgaattgcacacatttccctatagcttaggttt base pairs  
ctgtataaaaaaacagagcgacacctccctcacatgcatttgacttaacgtgtgtaaaaggaatatcgaaatccaaa 9376 to 9450

## SauIAI

ttgactgcgaactgttggtgcgtagctcaataactaaagtgtgatctggagcagctctacagtgaaatggttgaca base pairs  
aacgtacgttgacaaccacgtacacgcagttattgatttcaactagacctgcagatgtcacttattcaactcgt 9451 to 9525

ctgtgaaaatgtgcgtattttttcaaacgtcgccgactttttcctaatagaaaatgggcagtcaggcagtgat base pairs  
gaacattttacagctacataaaaaatgtttgcgaccgtgaaaaaaggattatcttttaccgcagtcgcgtcacta 9526 to 9600

## SauIAI

accgcttcgggtattttctcgtatgatgtggatctggagcgtcatctagtagtgcacagaggacagcaacccctt base pairs  
tggcgaaagccataaagaagactactacacctagacctctgcagtagatcatcgggtgtcctgtcgttgggaa 9601 to 9675

## SauIAI

cttgatctgtattctcgtggtggggcatgtctactggtctttgtctgggtgcagctctttctggcttgaaactt base pairs  
gaactagacataagaccgacaccccccgtacagatgaccagaagacgaacccacgtcgagaagaccgaactttgaa 9676 to 9750

## SauIAI

gaaactgttaacttctaactagtcattctgttgatagaatattgttcacgatctcagaacttattctattgtctgc base pairs  
ctttgacattagaagatcgtacgtagacaaacctctctatacaagtgctagagctcctgaaatagaataacaagac 9751 to 9825

## MseI

gcttgacagcagtggtgctgtgatttcaacagttttgcgtgccaaagttaaataatcatccacagactgagg base pairs  
cgaaactcgtctaccgcacaactaaagttgtcaaaacgcaggtttcaaattttatattagtaggtgtcgtactcc 9826 to 9900

## HinfI

tatggatagtaaaacttcatcttggattccatctgttctgcagctactcttcaaaagtgtcggatttttggtatg base pairs  
atacctatcatcttgacagtagaaccctaaggtagacagaacagtcgatgagaatgtttcacagacctaaaaacctac 9901 to 9975

## MseI

taggtgcgaatgagaaagccgatgagttcttcgcctccttaaggaatgggcagctctatgcacaaaatattgct base pairs  
atccagcgtttactcttctcggtactcaagaacgggaggaattcttaccctcgaagatagctttttatacag 9976 to 10050

## HinfI

## SauIAI

caagatgtggattcaaaatttagcaagctcagaggatcaagcggatgaagacagccctcttgcaaggacgaattt base pairs  
gtttctacacctaaagtttaaaactcttgaagctccttagctcgctactctctgtcgggagaactgttctcgttaa 10051 to 10125

## Seq2RM

## HindIII

## MseI

gttgtagagaagcttgcgcggatattgtatgggtgcagtcagcagggaaaaatggcattcttttaaggtacttcag base pairs  
caacatctcttcgaagcgcctatacaataccacgcgtcactgcctcctcttaccctagataaaaattccatgaagtc 10126 to 10200

## HinfI

## MseI

tgtcattttgtcattcttacttgatttccacacaaaaaatcaattacttaagccctgtcaaacatggatatttctgc base pairs  
acagtaaacaaagtaagatgaactaaaggtgtgttttttagttaatgaatttcggacagtttgcctacctataaagac 10201 to 10275

## PecI

## HaeIII

tatatatttgcgtgaacgtctagatttctgcaggtccagtggaaggatacggccctgaggaggatacatagggaacc base pairs  
atataaaacgacttgcgtctaaagacgtccaggtcaccttctctatgccgggacctctctatgacctctgg 10276 to 10350

## AvaII

gattgataactttgaggttagtgataggtatattcgtctgtgttgccttctgataacctatttgcattcaatccttg base pairs  
ctaactattgaactccaatcacataccatataagcagacgaacacacacataatgataaaacgttagattggaac 10351 to 10425

## Seq1RM

## HinfI



FIG. 5

Continued

ttttgcaaacagtgactgccgcagaaaaatttagagaatttgtacaagaaggggcacaaaagaagattctccac base pairs  
aaaacgcttggctcactgaacggcgctctttaaactctctaataacatgttcttcccgctgttcttcttcttaagaaggggtg 10426 to 10500

S3eR

EcoRV  
 TGCCTGTGAGTATTAGTCGTTGTGATTTGCTCGCTATTTGTTAGTCCCTTTTATTGGTGATATCTG base pairs  
 ACGGCACTCATAAATCAAGCAACACTAAAACGAGCGATAACAATCGAGGGGAAAAATAAACCATATAGAC 10501 to 10575

**EcoRV**

HaeIII
MseI

cctattttattctttcaagggtgatgttgatgtcatttgtggaggccaccatgccaaggatcagtggtttaa base pairs  
ggataaaataagaaatttccactacaactacagtaaacacctcgggtggtacggttcatagtcacccaaat 10576 to 10650

HaeIII

MseI

3F

atcgggtacagaaacggtgatgacacctcaaaagatgagaaaaacaaacaaatggtagcttccatggatattgtgg base pairs  
tagccatgtctttggcactactcggtaggttccactctcttttgtttgtttaccactgaaagtacctataacacc 10651 to 10725

3R

cgctacttgaagcccaagtatgttctcatggaaaaatgtgggtggacatactcaaatttcgggatggttacctaggaa base pairs  
gcacgaactctcgggttcatacaagaatcaccctttcacaccactctatgaatttaaacgccctaccaatggatcct 10726 to 10800

Hba I

IF

aatatgctttagctgcttctgtgctatgaagtacaaagcgcggcttggaaatgatggtggcgggttctatgggc base pairs  
ctatcgaacactcgacggaacaacgatactcatggttcgcgcgcgaaccttactaccaccgaccaacgataccag 10801 to 10875

Page 2A

tgccacagttcaggatgcgtgtgtgacctctgggggtctcttctctctgggtctgtctctgtacctgtctgttta base pairs  
acqgtgtcaagtctctacgcacacatgccagaccccccacagagaaagagcttcacagacaagcatggaacgacaaaat 10876 to 10950

Race 2B

ScrFI

ECORII

HinfI EcoRII  
 tatgtcttcctgctagatctcatattgctactgttggtctgctggctaaccagggtgtacgtgtatttgcaaatattagctgc base pairs  
 ataccgaacgcgtctcaactatataacgtcacaaacccacacacccatctgtccacatgcacataaacttttaaatccacg 10951 to 11025

BstNI

Race1A                  4F  
tctccatggtcttcctcgcgcccaactagacattgttttagtacgtcgggaggagcccctaatagccttttcggtgagtgcgaat base pairs  
ccccatcatctatggagacacggttttgtaactacacaacatcatcacctctctcggqattacgqaagaaccactcacatta 11026 to 11100

4F

Race1A

Race1B

**RaceRT**

cacaacacactactatgaaatcatgtggaatgtgaaatacgtgaccaaactgaattgttgagcaaatgtatg base pairs  
gtgtttgtgtatgatacttttagcaccttacacattttatgcgactgttgacttaacaacgtctgacacac 1101 to 1175  
  
ctgtcatatgacgacgacacaaaacacacgttgaataaactgtctttctggcgatgcatttccagatttacc base pairs  
gaagctcatctgctctgttttttgaacacacattttctggacgaacaaacagctctaaattgcaattgtac 1176 to 1250

MseI

PstI

aaaggcaagtgttctgtcaagttcatgcatcttctcagtgcagcatgtatttaactcttctctgcaggttcaaaatc base pairs  
ttctcgttcacaaacagttcgaagtcacgtacgtataaattgagaagacgtccaagtttag 11251 to 11325

ECORI

HhaI

Tagi

accagcctaacgatgtgatggagatgtgtgggtcccccaagaccgaattccagcgctacattcgactcagtcgtaa base pairs  
tggtcggattgctacactacatcataccaccaaggggttctggtctaaggctcgcgatgtaagctgagtcagcatt 11326 to 11401

Hinf I

HaeIII

**MaeI**

aggtaaaaaaccccggaactactactggttgcccttactacgaatatgttaggatttaatttcagaagaaccg base pairs  
tccattttttggggcacttgatgatqaccaaccggaagtgatgcttatcaattcctaaattaaagtcttcttggc 11401 to 11475

Part I

Ava II

ccctttttttcttggtgcttcggtactactgcagcaagctcactcttattatcatgtcagacatgttggttggt base pairs  
ggaaaaaaaagaaccacgaagccatgatgacatcttcgaagtgaagaataatagatcacgtctctacaacctaacca 11476 to 11551

18/39  
FIG. 5

AvaII Continued Sau3AI MaeI

cccttcggtgaagggggtggtccagatgaaggcgaagcctcttgatcaccagcctttacgcttaacacgatgat base pairs  
ggaagccacttccccgcaggctcacttccgttgcagaaacctagtggtcggaatgcggaattgttgcacta 11551 to 11625

HinfI

catgagcgggtcacaacgatctcctgccaagaaggttggtggttgcgcatttgccttcccttggttgtttt base pairs  
atattcgccaagttgtcttaaggacagctcttccaaccacggaaccgcgtaaacacggaaggaacaacaaaa 11626 to 11700

5F

tccccctctgaacaatcatctcttctccatgacaggagccaacttcgcgcacctaagggcctgaagggttg base pairs  
agggaagaccttgttagtagagagaaggatagctccctcggttgaaggcgctggatttccgcactcccaac 11701 to 11775

BamHI TaqI

gagcaaacatatttggtaggggatccagaagatcgagctgtgaactttcatctgggaaaccactggtatgtg base pairs  
ctggtttgtataacaactcaacctaggcttttagctgcacactttgaaagtacacccctttggcgaccatacac 11776 to 11850

Sau3AI

tgctattccgctgctgtgttctctataactgtgcaacatttcttcccatttcaaacctataactgacgaga base pairs  
acgataaaaggcgcgaacaaaggatattgacacgttgtaaatgaaagggtataaagtttagattgactgctct 11851 to 11925

HinfI

tgctgcacactactgtgaagattcatggctaacccatgacaaacattttgcacacatcttggttatctaggttccctga base pairs  
acgagcttgatgacattctaatgaccgattgggtactgtgttaaacgtgtgtgagaacaatagatccaaggact 11926 to 12000

ctatgcgaattcatctacaaaggccasatcactcaagtaagttcaaaacatttttttgggtttggggaaaa base pairs  
gatacgttacagtaagtagtcccggttagtgagttcattcaaaagtttggtaaaaaaaacaaaaacccctttt 12001 to 12075

HaeIII HhaI

gtaggttattgttacttgtgcttcatatgatgtgtgcaggccggttggggcgcctgtggtgggacgaagacgttc base pairs  
catccaataacaaatgaacacgaatgtatactacaagctccggcaaacccgcggacaccacccctgttctgtcaag 12076 to 12150

ScrFI

HaeIII

ctacagttgtaaccagagcagagcctcacacagggtcagcttcagaaggccactccttttcgcgaatccctgc base pairs  
gatgtcaacattggctctcgtctcggagttgtgtccagtcgaagctcttccggtgaggaaaaacggttagggagc 12151 to 12225

BstNI

Sau3AI

atctgtatttactattagtcgtgtgttccatgatgatcattaccgaacattgtgtccacacaggttataattatc base pairs  
tagacataaatgataatcgacacagaaggtatactagtaatggttgtacacagggtgtgtccaatattaaagtag 12226 to 12300

ScrFI

HinfI

Rco0109I

HpaII

cgactcaagcaagggtccctcactatccgggagagaacgcaaggttacagggttccccgattattaccgattgttg base pairs  
gctgagctcgttccmaggagtgatagggcctcttgcgttccaatgtccccaagggtcataatggtaacacac 12301 to 12375

AvaII

MspI

HaeIII

Sau3AI

gctcgatcaaggaagaagtaagttccgtgttttcaagttgctgtaccagatctagtcactattgaaagttttcagc base pairs  
cgggttagtttctcttattcaaggacaaaaggttcaacggcatggtcctagatcagtgataactttcaaaagtcg 12376 to 12450

Sau3AI

BglII

agcaagccattcatcagtagttacagctcttgaagccttacctctgaacattgtgtcttctctgaggtgat base pairs  
tcggttcgtagaattgatcaatcaatgtcgagaacttccgaatggagactgtacacacgaaagagactaccacta 12451 to 12525

MspI

HpaII

EcoRII

aggatcattcaagtcgggaacgcagtggtgtcctgttcccgggcactgggtactgtctggggcaagccac base pairs  
tccatgttaagttcagcccttgcgtcaccgcaggggacaaaggcccggtgacccgatgacagaccccggtcggatg 12526 to 12600

ScrFI

SmaI

19/39

FIG. 5

ScrFI

HinfI

PvuII

Continued

ctgggtgaatctgaagggagtgacccctctgtaccagctgctccaagtttcacctctgttggaggacgcactgcg base pairs  
 gaccacttagactccccctcactgggagacatggtcgacggaggttcaaaagcggagacaacctctctgctgacgc 12601 to 12675  
 BstNI

Eco0109I

PstI

Sau3AI

gggcaggcagggcctctcttctgtttggcacccttcgacggggaggttagttgacgagtaaaaggatgacagatctga base pairs  
 ccgctccgctcccggagaaggacaacctgtggggacgtccccctccatcaactcgtcatlctcttactgcttagact 12676 to 12750  
 HaeIII LR BglII

TaqI

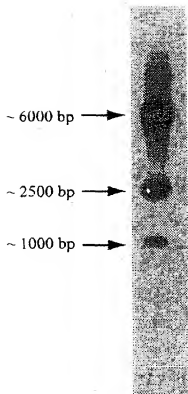
gctgagctgggcaacatccagcggcaggagcatttctggttcggttcgattcgggctcacga base pairs  
 cgactcgaccctgtgtaggtccgctcctcgtaaagaccaagccaagcctaagcccgagtgct 12751 to 12812  
 HinfI

FIG. 6

PROCESS	WORLD WIDE WEB SITE
sequence format conversion	<a href="http://dot.imgen.bcm.tmc.edu:9331/seq-util/Options/readseq.html">http://dot.imgen.bcm.tmc.edu:9331/seq-util/Options/readseq.html</a>
reverse complementation	<a href="http://dot.imgen.bcm.tmc.edu:9331/seq-util/Options/revcomp.html">http://dot.imgen.bcm.tmc.edu:9331/seq-util/Options/revcomp.html</a>
sequence translation	<a href="http://dot.imgen.bcm.tmc.edu:9331/seq-util/Options/sixframe.html">http://dot.imgen.bcm.tmc.edu:9331/seq-util/Options/sixframe.html</a>
protein information	<a href="http://www.expasy.ch/tools">http://www.expasy.ch/tools</a>
sequence alignments using Clustal W	<a href="http://dot.imgen.bcm.tmc.edu:9331/multi-align/Options/clustalw.html">http://dot.imgen.bcm.tmc.edu:9331/multi-align/Options/clustalw.html</a>
sequence comparisons using BLAST 2.0	<a href="http://www.ncbi.nlm.nih.gov/gorf/bl2.html">http://www.ncbi.nlm.nih.gov/gorf/bl2.html</a>
sequence searches using BLAST 2.0	<a href="http://www.ncbi.nlm.nih.gov/blast/blast.cgi?Jform=0">http://www.ncbi.nlm.nih.gov/blast/blast.cgi?Jform=0</a>

21/39

FIG. 7



**SUBSTITUTE SHEET (RULE 26)**

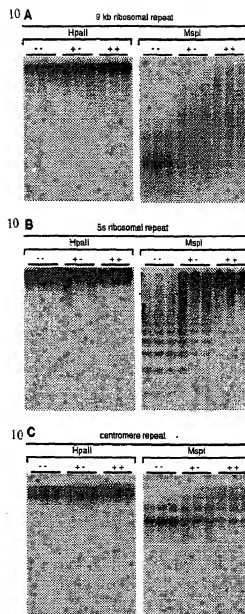
23/39

FIG. 9

Motif	SAM binding		Cytosine binding	
	<i>M.Hhal</i>	<i>zmet2a</i>	<i>M.Hhal</i>	<i>zmet2a</i>
I	Phe18	Try347		
II	Glu40	Gln407		
	Trp41	Trp408		
III	Asp60	Asp428		
IV	Pro80	Pro516	Phe79	Pro515
	Gln82	Gln82	Cys81	Cys517
V	Leu100	Val542		
VI			Glu119	Glu559
			Asn120	Asn560
			Val121	Val561
VIII			Arg165	Arg605
X	Asn304	Asn851		

24/39

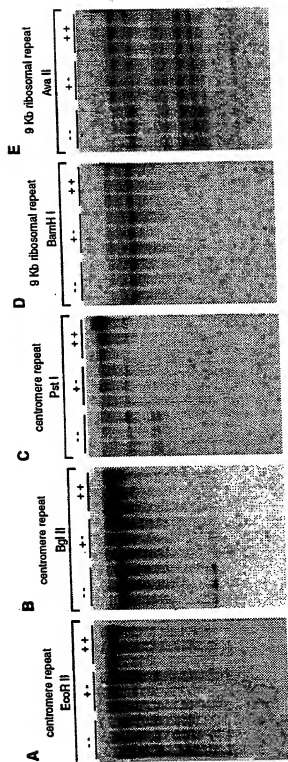
FIG. 10





25/39

FIG. 11



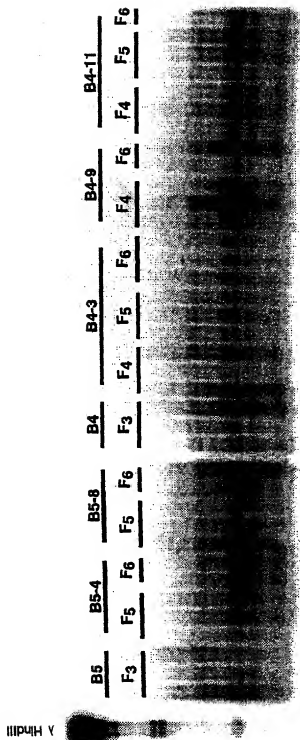
26/39

FIG. 12

GENOTYPE	NUMBER OF PLANTS	TOTAL 5mCytosine (%)	% WT levels	% decrease
wild type	3	34.40 $\pm$ 0.55	100	0.0
heterozygous zmet2a-mu1	7	32.00 $\pm$ 0.90	93.0	7.0
homozygous zmet2a-mu1	5	30.40 $\pm$ 0.19	88.4	11.6

2000-03-01 14:00:00

FIG. 13



28/39

FIG. 14

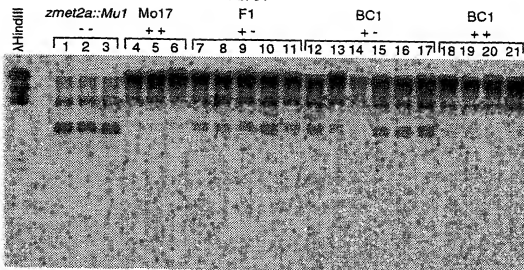
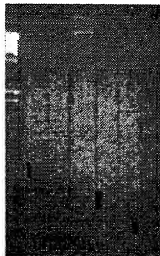


FIG. 15

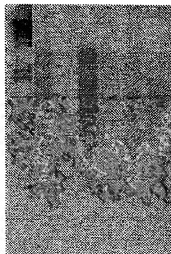
15 A

$\lambda$  HindIII  
 embryo 24 DAP  
 young leaf  
 immature ear  
 immature tassel  
 BMS callus  
 10 day seedling



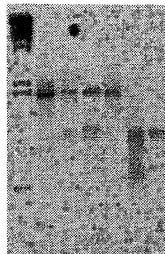
15 B

$\lambda$  HindIII  
 embryo 24 DAP  
 young leaf  
 immature ear  
 immature tassel  
 BMS callus  
 10 day seedling



15 C

$\lambda$  HindIII  
 embryo 24 DAP  
 young leaf  
 immature ear  
 immature tassel  
 BMS callus  
 10 day seedling



30/39

FIG. 16

5' LTR

catgc**TGTT**GGGGCCATGTGTCTAGTGTGGCCCATTAACGTGTACA  
CATATACTAGAAAGTGTGTGTGGTGTAGAGAGAGTGTCTATGTTTT  
CCACATTCCAGAAAAATCC**CACA**TGGTATCAGAGCCAGG

PBS

3' LTR

PPT

GAGGGGGAG**TGTT**GGGGCCATGTGTCTAGTGTGGCCCATTAACGTG  
TACACATATACTAGAAAGTGTGTGTGGTGTAGAGAGAGTGTCTGTATG  
TTTCCACATTCCAGAAAAATCC**CACA**catgc

31/39  
FIG. 17

	<b>Gag</b>	<b>Protease</b>
SPRITE-1	- CYNCGNVGHIARNC	TQVTOLKWLDSGASKH
hopscotch	- CQVCSRVGHTALNC	QNGSNVWYTDGTADH
retrofit	- CQVCFKRGHATAAC	SYGIDTNWYIDTGATDH
arabpolprt	- CSNCGRTGHEKKCC	GKTLKGLDILDSGASH
copia	- CHHCGREGHIKKDC	SVMDCNCGFLVDSGASH

**Integrase**

SPRITE-1	- QVKILRPDN-CTEYVNKGFNAFLSRNGILHQTSCPDTPPONGVAERKNRHILE
hopscotch	- KIIAFQSDW-CGE--YEKLNAHEKTTIGIHHQVSCPHTHQONGAERKRRHIVE
retrofit	- KIIAMQTDWRGGR--YQKLNSFFAQIGLIIMCHVLTILIRQNGSAERKRRHIVE
arabpolprt	- TVKMVRSDN-CTE--FMCLSSYFRENGI IHQTSVCGTTPQONGRVERKRRHILN
copia	- KVVVLYIDN-GREYLSNEMRQFCVKKGISYHLTVPHPTQLNGVSRMIRTI TE

**Reverse Transcriptase**

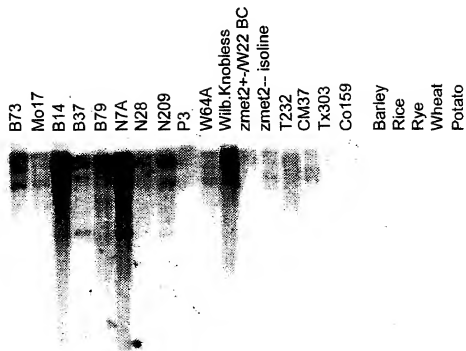
SPRITE-1	- RYKAALVARGYSQYTGIDYDETFAPVAMSTVRLISCAATFNRLYOLQVKNALHGDQGEVMEIPFG (59) AILAVTVDDIII
hopscotch	- RYKAALVARGFQYQYTGIDYDTPSPVVKHSTIRLDVLSAVSQWLSRQLQVQNAFLHGILILETVVMKQPPG (59) IYILTVDDIII
retrofit	- RYKAALVARGFQYQYTGIDYDTPSPVVKATIRIILSIASRQWLSRQLQVQNAFLHGFLLEEIVMQPPG (59) HFLVTVDDIIV
arabpolprt	- RYKAALVARGQVGESEYKSTFAPVVRMTVTYRLANVAANQWVYQMDVKNALHGDGLEEVYMKLPFG (59) LRVLTIVDDLLI
copia	- RYKAALVARGFTQYQYTGIDYDETFAPVARISSPFIILSLVTQYKLAVHQMVKTAFLNGTLKEEYMRLPQG (59) IYVLTIVDDVVI

**RNase H**

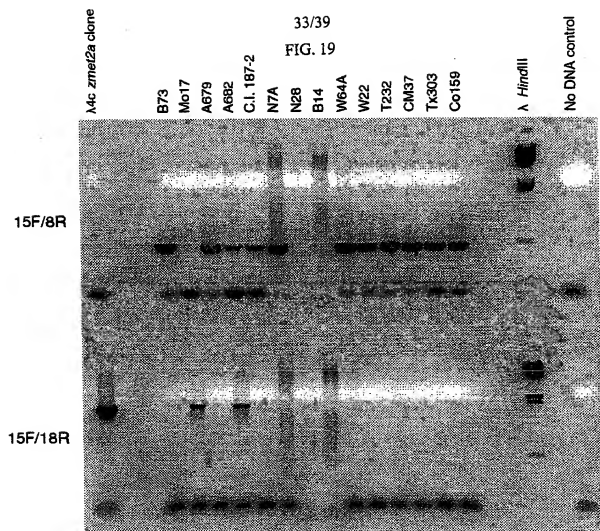
SPRITE-1	- DADWAGSLDRRSTSGYCVFVGG-NLVSWSKQSTVSRSSTAEYKANAIAICDMLIKGLL (25) NPVQHQRTKRIEVDHFF
hopscotch	- DADWAGSLDRRSTSGYAVFLGP-NLISWSKQSTVSRSSTAEYKANAIAIAEVLMLQSL (25) KPFFNARTKHIEVDHFF
retrofit	- DADWAGSLDRRSTSGYAVFLGP-NLVSWSKQSTVSRSSTAEYKAVANTAEVLMLVQTL (25) NPVFNARTKHIEVDHFF
arabpolprt	- DSDWAGSLDRRSTSGYAVVLLGG-SPIWSKTRQSTVSRSSTAEYKANSYALKEIKMLRKL (25) NPVFNARTKHIESDCHS
copia	- DSDWAGSEIDRSTSGYLVNMFDFNLICWNTKQNSVAASSTAEYKALFEACREALMLKFL (25) NPSCNARKHIDTKYH

32/39

FIG. 18

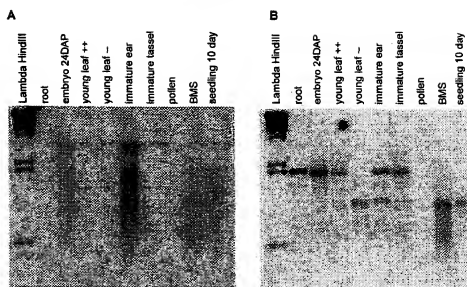


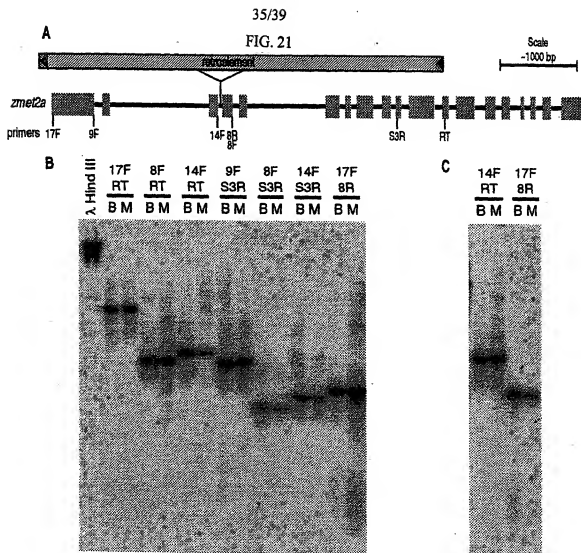




34/39

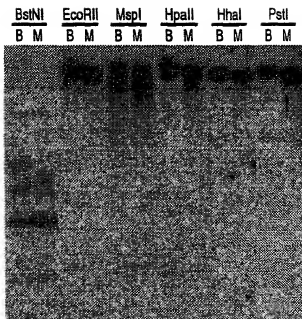
FIG. 20





36/39

FIG. 22



37/39

FIG. 23

GGGAATTCGATTACTCACTATAGCGCTCGAGCGGCCGCCGGGCGAGGTTGCGAAAACCATC  
AACCTAACGATGTAATGGAGTATGGTGGTTCCCCCAAGACAGAGTTCCAGCGCTACATTC  
GACTTGGTCGTAAGACATGTTGGATTGGTCGTTTGGTGAGGAGGCTGGTCCAGATGAAG  
GCAAGCTCTTGGATCACCAGCCCTTACGGCTTAACAATGATGATTAGAGCGGGTTAAGC  
AAATTCCGTGCAAGAAGGGAGCCAACTTCGTGACCTAAAGGGTGCAAGGTTGGAGCAA  
ATAATGTTGTTGAGTGGGATCCAGAAGTCGAACGTGTGTACCTTTCGTCTGGGAAACCCAC  
TGGTTCCTGACTATGCGATGTCTATTCATCAAGGGCAAACTCACTCAAGCCATTCCGGCGCC  
TGTTGGTGGGACCAAGACGGTTCTACAGTTGTGACCAGAGCAGAGCCTCATAACCAGGTTAT  
ATTGCATCCGACTCAAGCAAGAGTCTTGACTATCCGGGAGAACGCCAAGGTTACAGGGCTT  
CCCCGATTACTACCGATTGTTTGGACCGATCAAGGAGAAGTATATTCAAGTCGGGAACGC  
AGTGGCAGTCCCTGTTGCACGGGCACTGGGCTACTGTCTGGGTCAAGCCTACCTGGGTGA  
ATCTGACGGGAGTCAGCCTCTGTACAGCTGCCTGCAAGTTTACCTCTGTGGGGCGAAC  
CGCGGTTCAAGGCAATGCCGCTTCTGTTGGCACTCCTCGGGGGAGGTAGTCAGGCAGTA  
AAAGGATAGCGGAGCAACCCCTGGTTGGTATTTTGATTGAGGCCATCCAGTAGCATGTTT  
ACCAATAAATAATCATTGGTCTGTCTGATTCTTATGGTTGGAAGATGAATGTATGTAGGGT  
GTACTCGAGCTCGAGTGCTTGTGTTGTAAGTTGAGGTTTCTCATCCATTGGCCTGCC  
TATTTGTGGATGACGTTTCAATTCAGATTAGCAATGTGCTTATTTAAGGTTTCGTCAATG  
ACCTGTATTCTACAATCCACTATTGTTTCCAAAGACAGCATTTGATCCTTAAAAAACT  
GTAAAAAAACAGTGCCTGAAAGCGCAAAAAACCTGCCC  
GGGCGGCCGCTCGAGCCCTATAGTGAGTAATCGAATTCCT

38/39

FIG. 24

EFDYSL\*RSSGRPGRFENHQPNQVMEYGGSPKTEFQRYIRLGRKMDLQWS  
FGEEAGPDEGKLLDHQPLRLNDDYERVKQIPVKKGANFRDLKGVKVGAN  
NVVEWDPEVERVYLSSGKPLVPDYAMSFYKSLKPFGRWWDQTVPTVV  
TRAEPHNQVILHPTQARVLTIRENARLQGFDPDYRFLFGPIKEKYIQVGNA  
VAVPVARALGYCLGQAYLGESDGSQPLYQLPASFTSVGRITAVQANAASVG  
TPAGEVVEQ\*

100  
90  
80  
70  
60  
50  
40  
30  
20  
10  
0  
100  
90  
80  
70  
60  
50  
40  
30  
20  
10  
0

39/39

## FIG. 25

667 KVNQHQPNDVMEYGGSPKTEFQRYIRLSRKDMLDWSFGEGAGPDEGKLLDHQPLRLNND 726  
+ +NHQPNDVMEYGGSPKTEFQRYIRL RKDMLDWSFGE AGPDEGKLLDHQPLRLNND  
15 RFENHQPNDVMEYGGSPKTEFQRYIRLGRKDMLDWSFGEEAGPDEGKLLDHQPLRLNND 74

727 YERVQQIPVKKGANFRDLKGV RVGANNIVEWDPEIERVKLSSGKPLVPDYAMSFYKGS 786  
YERV+QIPVKKGANFRDLKGV+VGANN+VEWDPE+ERV LSSGKPLVPDYAMSFYKGS  
75 YERVKQIPVKKGANFRDLKGVKVGANNVVEWDPEVERVYLSSGKPLVPDYAMSFYKGS 134

787 KPFGRLWWDETVPTVVTRAEPHNQVIIHPTQARVLTIRENARLQGFPDYRFLFGPIKEY 846  
KPFGRLWW+TVPTVVTRAEPHNQVI+HPTQARVLTIRENARLQGFPDYRFLFGPIKEY  
135 KPFGRLWW+TVPTVVTRAEPHNQVILHPTQARVLTIRENARLQGFPDYRFLFGPIKEY 194

847 IQVGNAVAVPVARALGYCLGQAYLGESEGSPLYQLPPSFTSVGGRTAGQARASVPVGT 906  
IQVGNAVAVPVARALGYCLGQAYLGE+GS PLYQLP SFTSV GRTA QA A+ VGT 847  
195 IQVGNAVAVPVARALGYCLGQAYLGESEGSPLYQLPASFTSV-GRTAVQANAASVGT 253

907 GEVVEQ 912  
GEVVEQ  
254 GEVVEQ 259